A study on Face Detection in Images

Rajesh Kumar Mahapatra (110cs0132)

Department of Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela - 769008, India

A study on Face Detection in Images

Thesis submitted in

May 2014

to the department of

Computer Science and Engineering of

National Institute of Technology Rourkela

in partial fulfillment of the requirements

for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

Rajesh Kumar Mahapatra

[Roll : 110cs0132]

Under the guidance of

Prof. Pankaj Kumar Sa

Computer Science and Engineering

National Institute of Technology Rourkela

Rourkela-769 008, India. www.nitrkl.ac.in

Certificate

This is to certify that the work in the project entitled A study on face detection in images by Rajesh Kumar Mahapatra is a record of his work under my supervision in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering.

Dr. Pankaj K Sa

Acknowledgment

I take this opportunity to express my profound gratitude and deep regards to my guide Prof Pankaj Sa sir for his exemplary guidance, monitoring and constant encouragement throughout the course of this project. The blessings, help and guidance given by him time to time shall carry me a long way in the journey of life on which I am about to embark.

I am obliged to all the professors of the Department of Computer Science and Engineering, NIT Rourkela for instilling in me the basic knowledge about the field that greatly benefitted me while carrying out the project and achieving the goal.

Finally, I dedicate the thesis to my parents for their love, support and encouragement without which this would not have been possible.

Contents

List of Figures

Abstract :

Face detection has been one of the most active research topics in computer vision over the past decade. Viola-Jones face locator has accepted an extraordinary consideration, and has turned into the defacto standard of face identification approaches. In this project, we have used the Haar-like features to detect faces. Only the frontal faces can be detected. The face detection is done in three phases with the algorithm specified by Paul Viola and Michael J. Jones . We get some of non- faces which are classified as faces along with the faces. We tried to reduce those non-faces from the image and keep only faces using two algorithms. The first algorithm is based on observation method of the errors and removing the non faces based on a certain threshold . The second method detect the upper body and then keep the faces which are present in the upper body and discard the non faces which are not present inside the upper body region. By using the two algorithms, we tried to minimize the non-faces in the algorithm . We have taken real life examples and simulated the algorithms in MATLAB successfully.

Chapter 1

Introduction

1.1 Face Detection

Computer vision, when all is said in done, means to double (or in a few cases compensate)human vision, and customarily, have been utilized as a part of performing routine monotonous undertakings, for example, classification in monstrous mechanical production systems. Today, scrutinize on machine vision is spreading gigantically so it is very nearly difficult to organize every last bit of its subtopics. Notwithstanding of this, one can rundown important a few provisions, for example, face processing (i.e. gesture recognition and face expresssion), machine human cooperation, swarm reconnaissance, and substance-based picture recovery. All of the applications stated above require detection of face, which can be simply viewed as a preprocessing step for obtaining the “object”. The face is our primary center of consideration in social life assuming an imperative part in passing on feelings and character. We can perceive various appearances adapted all around our lifespan and distinguish faces considerably after numerous years of division. This aptitude is very hearty in spite of numerous varieties in visual jolt because of maturing, changing condition and preoccupations, for example, glasses, facial hair, or changes in hairstyle.

Face detection is a technology that determines the sizes and locations of human faces in digital images. It recognize faces and ignores anything else, such as trees, bodies and buildings. Face detection might be recognized as a more general instance of face confinement. It is the center of all facial analysis, e.g., face localization, face recognition, face authentication, facial feature detection, face tracking, and facial expression recognition. Additionally, it is an essential strategy for all different requisitions, for example, feature conferencing, substance-based picture recovery and adroit human machine cooperation (HCI). The objective of face detection is to figure out if or not there exist any appearances in the picture and, if present, give back where its due and the degree of each one face. While face detection is an inconsequential assignment for human vision, it is a test for machine vision because of the varieties in scale, area, introduction, posture, facial articulation, light condition, and different appearance characteristics. Face discovery is utilized within numerous places now a days particularly the sites facilitating pictures like photobucket, picassa and facebook. The consequently labeling characteristic adds another measurement to imparting pictures around the individuals who are in the picture and likewise gives the thought to other individuals about who the individual is in the picture.

There exist two different types of face detection problems:

1) Face detection in images

2) Real-time face detection

1.1.1 Face detection in images

Most face detection frameworks endeavors to concentrate a small amount of the entire face, by dispensing with the majority of the foundation and different zones of a singular's head, for example, hair that are not important for the face distinguishment . With static pictures, this is regularly done by running a sliding "window" over the picture. The face detection framework then finds if a face is available inside the window. Lamentably, with static pictures there is a vast hunt space of conceivable areas of a face in a picture. Faces could be expansive or little and be situated anyplace from the upper left to the easier right of the picture.

1.1.2 Real-time face detection

Continuous-face recognition includes discovery of face from an arrangement of casings from a feature-catching gadget. Real time face detection is really a far more straightforward procedure than recognizing a face in static pictures. It is on the grounds that not at all like a large portion of our nature, individuals dependably keep moving. We stroll around, wriggle, wave our hands about, squint and so on.

1.2 Applications Of Face Detection

- **Facial recognition**

  Face detection is utilized within biometrics, as a piece of or together with a facial distinguishment framework. It is likewise utilized as a part of human machine interface, video surveillance and image database management.

- **Photography**

Some recent digital cameras use face detection technique for autofocus. Face detection is useful for selecting regions of interest in photo slideshows that use a pan-and-scale Ken Burns effect.

- **Marketing**

Face detection is picking up the enthusiasm of all advertisers. A webcam might be coordinated into a TV and identify any face that strolls by. At that point the framework computes the sexual orientation, race and age extent of the face. Once the data is accumulated, an arrangement of notices might be played that is particular towards the identified age/race/sex.
An example of such a framework is known as Optimeyes and is coordinated into Amscreen digital signage system.

- **Smart captcha**

It is a mixture of an effectively existing captcha which utilizes sounds and realistic pictures . On the other hand, utilizing a face or movement discovery engineering we are not going to trouble clients with perceiving peculiar letters and vague sounds any more. All that the client needs is to show his face in movement so that the site holder could make certain that he is an individual, and not a machine.

1.3 Different Face detection techniques

- Appearance-based and learning based approaches
- Find faces by color
- Find faces by motion
- Find faces by using a mixture of color and motion
- Find faces by Neural network and kernel based approach
- Find faces by weak classifier cascades

1.4 Literature Review

1.4.1 Face detection

Early exertions in face detection have gone over as promptly as start of the 1970s, where basic anthropometric and heuristic systems were utilized. These systems are generally unbending due to different presumptions, for example, plain foundation, frontal face a common visa photo situation. To all these frameworks, any change in picture conditions might mean a fine-tuning, if not a complete overhaul. Regardless of all these issues, the development of exploration investment stayed steady until the 1990s, when handy and genuine face distinguishment and feature coding frameworks began to turn into an actuality. Over the past few decades there has been a lot of examination enthusiasm traversing different critical parts of face identification. More hearty division plans have been introduced, for the most part those utilizing color, movement and summed up data. The utilization of neural systems and facts has likewise empowered appearances to be identified from jumbled scenes at different separations from the Polaroid. Moreover, there are different developments in the configuration of characteristic extractors, for example, deformable layouts and the dynamic shapes which can find and track the facial characteristics appropriately. The ash data inside a face can additionally be utilized as characteristics. Facial characteristics, for example, understudies, eyebrows and lips show up for the most part darker than their encompassing facial locales. This property could be misused to separate different facial parts. Different late facial characteristic extraction calculations hunt down nearby light black minima inside divided facial regions. In these calculations, the info pictures are first upgraded by complexity-extending and ash-scale morphological schedules to expand the nature of neighborhood dim patches and in this way make location less demanding. The extraction of dim patches is accomplished by low-level ash-scale thresholding. On the provision side, Wong et al. execute a robot that searches for dim facial districts inside face applicants got in a roundabout way from shade examination. The calculation makes utilization of a weighted human eye layout to focus conceivable areas of an eye pair. In Hoogenboom and Lew, neighborhood maxima, that are defined by a brilliant pixel encompassed by eight dull neighbors, are utilized rather to show the splendid facial spots, for example, nose tips. The discovery focuses are then adjusted to the characteristic formats for connection estimations. Yang and Huang, then again, investigated the light black-scale conduct of countenances in mosaic (pyramid) pictures. At the point when the determination of a face picture is diminished either by averaging or subsampling , naturally visible characteristics of the face will vanish. At low determination, face locales will get uniform. Taking into account this perception, Yang proposed a various leveled face discovery schema. Beginning from low determination pictures, face hopefuls are dictated by a situated of decides that hunt down uniform areas. The face hopefuls are then verified by presence of conspicuous facial characteristics utilizing neighborhood minima at higher resolutions. The method of Yang and Huang was incorporated into a framework for rotation invariant face recognition by Lv et al. furthermore an expansion of the calculation is displayed in Kotropoulos and Pitas.

1.5 Motivation

The Motivation proposing the new algorithm is :-

- A better method which can detect the faces and eliminate non-faces is developed.

Chapter 2

Viola Jones Face Detection Method:

2.1 About the method

The fundamental rule of the Viola-Jones calculation is to sweep a sub-window fit for catching faces over a given info picture. The fundamental picture transforming methodology might be to rescale the information picture to different sizes and after that run the altered size locator through these pictures. This methodology ends up being tedious because of the figuring of the diverse size pictures. In opposition to the standard methodology Viola-Jones rescaled the detector rather than the data picture and run the locator commonly through the picture each one time with an alternate size. From the get go one may suspect both methodologies to be similarly tedious, however Viola-Jones have concocted a scale invariant detector that requires the same number of computations whatever the size. This locator is built utilizing a supposed-fundamental picture and some basic rectangular characteristics reminiscent of Haar wavelets. Viola-Jones face indicator has three recognized key commitments, which prompt high handling rate and discovery rates. These key commitments are: the basic picture, a proficient taking in calculation focused around Adaboost, and a course structure.

2.2 The scale invariant detector

The primary venture of the Viola-Jones face discovery calculation is to transform the info picture into a fundamental picture. It is carried out by making every pixel equivalent to the whole total of all pixels to the left or more of the concerned pixel. This takes into consideration the computation of the whole of all pixels inside any given rectangle utilizing just four qualities. All these qualities are the pixels in the indispensable picture that correspond with the corners of the rectangle in the info picture.

The Viola-Jones face detector dissects a given sub-window utilizing characteristics comprising of two or more rectangles. There are 5 of such characteristics. Each one characteristic brings about a solitary worth which is computed by subtracting the total of the white rectangle from the aggregate of the dark rectangle. Viola-Jones have emperically observed that a finder with a base determination of 24*24 pixels gives palatable effects. While considering all conceivable sizes and positions of the characteristics what added up to give or take 160,000 separate characteristics can then be built. Consequently, the measure of conceivable characteristics boundlessly dwarfs the 576 pixels held in the locator at base determination.

The 5 basic Haar-features are :

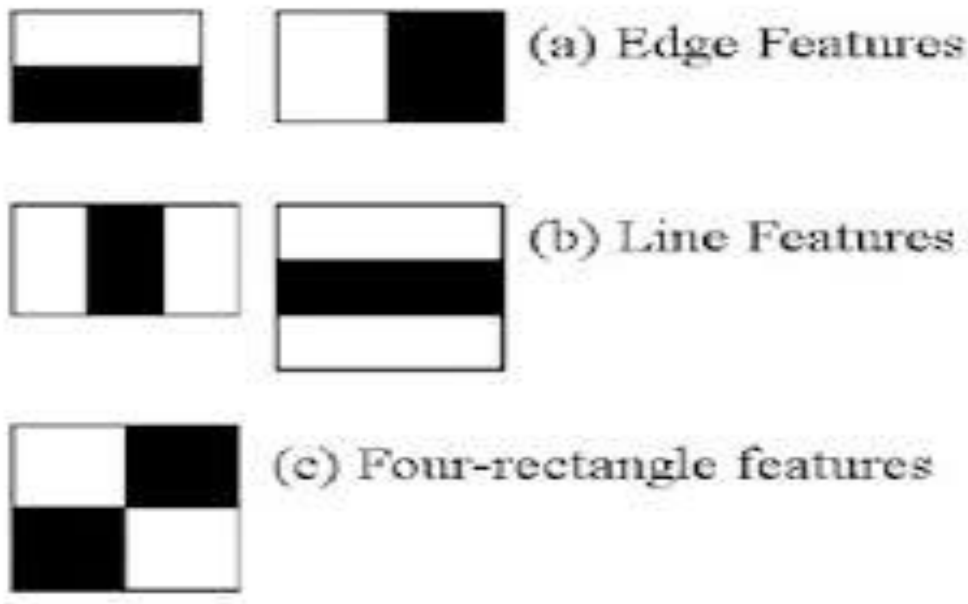

Figure – 2.1

Rectangle_Feature_value f = $\sum$ (pixels values in white area) – $\sum$ (pixels values in shaded area)

If Rectangle_Feature_value > threshold, then feature =1 else feature = 0

The integral image at location (x,y), is the sum of the pixel values above and to the left of (x,y) .

Input image :

1	1	**1**
1	1	1
1	1	1

Figure – 2.2

Integral image :

1	2	**3**
2	4	6
3	6	9

Figure – 2.3

If A,B,C,D are the values of the integral images at the corners of the rectangle R. The sum of image values inside R is:

Area_R = A + D – B – C

If A,B,C,D are found , only 3 additions are needed to find Area_R

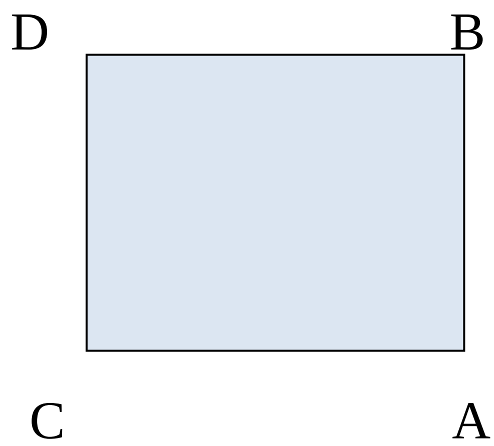

Figure – 2.4

2.3 The modified AdaBoost algorithm

As expressed above there might be computed more or less 160.000 characteristic values inside a finder at base determination. From all these characteristics some few are required to give very nearly reliably high values when on top of a face. To discover these characteristics Viola-Jones utilize a changed variant of the Adaboost calculation created by Freund and Schapire in 1996 . Adaboost is a machine taking in boosting calculation fit for developing a solid classifier through a weighted mix of feeble classifiers. (A frail classifier orders accurately in just a tiny bit more than a large portion of the cases.) To match this wording to the exhibited hypothesis each one characteristic is acknowledged to be a potential feeble classifier. A paramount some piece of the changed Adaboost calculation is the determination of the best characteristic, extremity and edge. There appears to be no brilliant answer for this issue and Viola-Jones recommend a straightforward beast power system. This implies that the determination of every new feeble classifier includes assessing each one characteristic on all the preparation cases keeping in mind the end goal to discover the best performing characteristic. This is required to be the most time intensive some piece of the preparation technique. Adaboost is a machine taking in boosting calculation equipped for developing a solid classifier through a weighted mixture of frail classifiers. To match this wording to the displayed hypothesis each one characteristic is acknowledged to be a potential feeble classifier.

A weak classifier is mathematically described as:

$$h(x,f,p,\theta) = \begin{cases} 1 & \text{if } pf(x) > p\theta \\ 0 & \text{otherwise} \end{cases}$$

where x is a 24*24 pixel sub-window, p the polarity, f is the applied feature and θ the threshold that decides whether x should be classified as a positive (a face) or a negative (a non-face).

2.4 The cascaded classifier

The essential standard of the Viola-Jones face detection calculation is to sweep the detector ordinarily through the same picture each one time with another size. Regardless of the fact that a picture ought to hold one or more faces it is evident that an unreasonable expansive measure of the assessed sub-windows might even now be negatives (non-faces). This acknowledgment prompts an alternate definition of the issue:

Instead of finding faces, the algorithm discard non-faces.

The thought behind this explanation is that it is quicker to dispose of a non-confront than to discover a face. In view of this an indicator comprising of one and only (solid) classifier abruptly appears wasteful since the assessment time is consistent regardless of the data. Thus the requirement for a fell classifier emerges. The fell classifier is made out of stages each one holding a solid classifier. The occupation of each one stage is to figure out if a given sub-window is certainly not a face or possibly a face. At the point when a sub-window is ordered to be a non-confront by a given stage it is promptly tossed. On the other hand a sub-window considered a possibly-face is passed on to the following stage in the course.

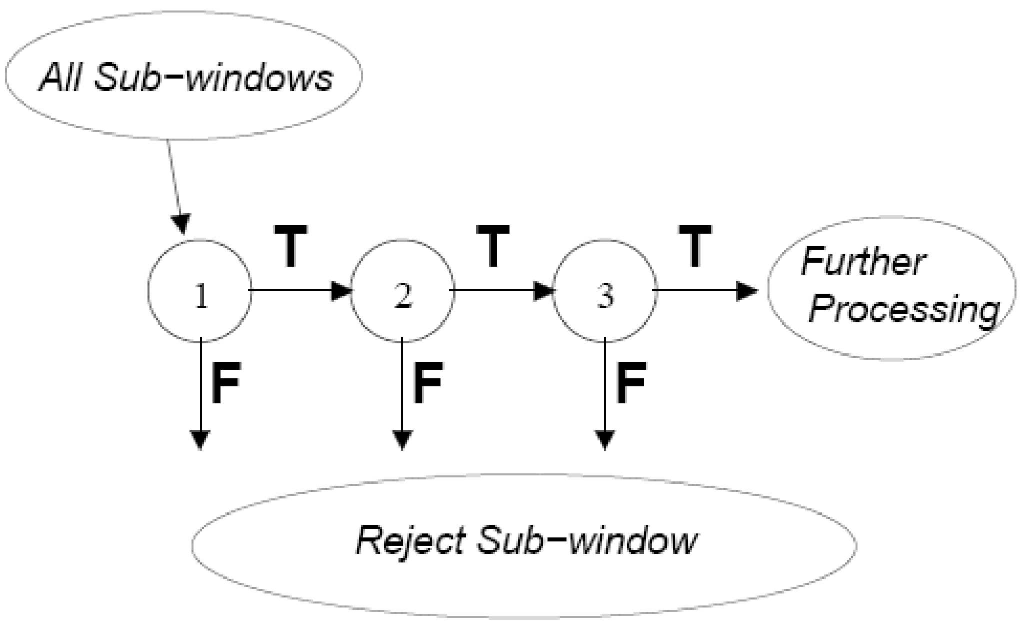

Figure – 2.5

2.5 Removing Non-faces:

We tried to implement 2 different methods to remove the false positives.

First method is given as follows:

- W = width of the box
- We find out sum of width of all boxes.
- We compute a threshold value which is average of the sum.
- For each box, if the w>threshold value then we store these values .
- Only these values are used to detect faces and other values are not considered.
- Then, we display the square boxes corresponding to those values.

The second method is described as follows:

- Detect all the faces using Viola Jones Face detection technique.
- Detect all upper body using Viola Jones method.
- Remove any false positives in upper body using average threshold method.
- Find all faces which are within the upper body region and discard those which are not in the region of upper body.

- If still any false positive remains, remove by using average threshold method.

By using both these methods, we reduce the rate of false positive (i.e, non-faces which are classified as faces) .

Chapter 3

Adaboost Algorithm Pseudo code

- Given examples images (x1,y1),(x2,y2),...,(xn,yn) where y1=0,1 for negative and positive examples.
- Initialize weights $w_{1,i}$ = 1/2m, 1/2l for y1=1,0 ,where m and l are the numbers of positive and negative examples.
- For t=1,2,3…..T :

1. Normalize the weights. $w_{t,i} = w_{t,i} / \Sigma_{j=1}^{n} w_{t,j}$
2. The error is evaluated as $\mathcal{E}_t = \Sigma_i w_i \, | h_t(x_i) - y_i |$
3. Choose the classifier h_t with lowest error $\mathcal{E}_t$.
4. Update the weights:

   $w_{t+1,i} = w_{t,i}\beta^{1-e_i}$

   where $e_i = 0$ if x_i is classified correctly else 1.

   and $\beta_t = \mathcal{E}_t / (1- \mathcal{E}_t)$

5. The final strong classifier is:

$$C(x) = \begin{cases} 1 & \text{if } \Sigma_{t=1}^{T} \alpha_t \, h_t(x) >= \frac{1}{2} \Sigma_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases} \qquad (\text{where } \alpha_t = 1/\beta_t)$$

Chapter 4

Results

Using 1st method :

Figure 4.1 Sample Image

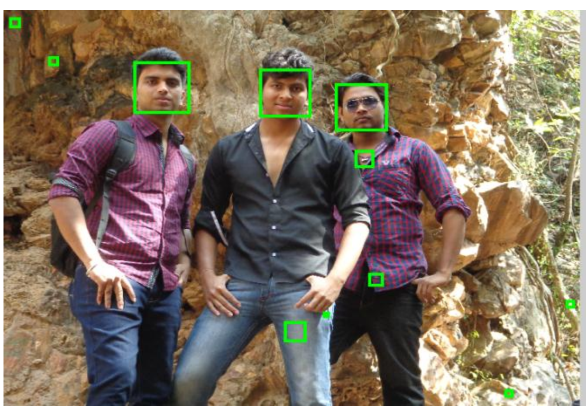

Figure 4.2 Face detection

Figure 4.3 Removing non-faces

Using 2nd Method :

Figure 4.4 Sample Image

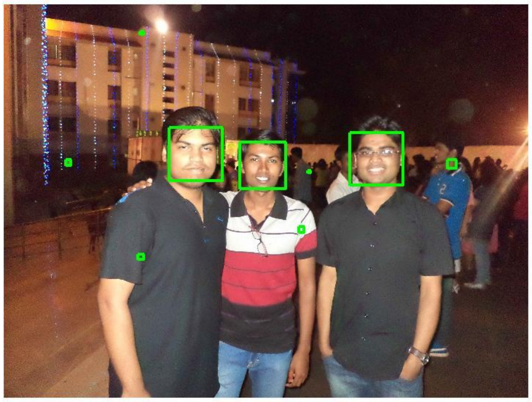

Figure 4.5 Detection of faces

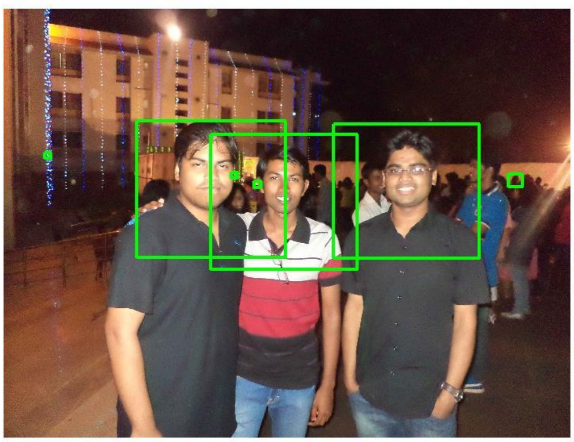

Figure 4.6 Detection of upper body

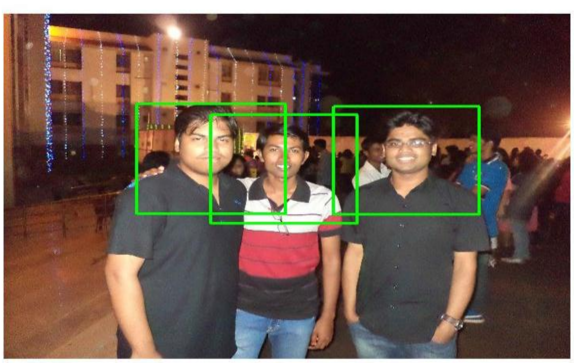

Figure 4.7 Removing non upper body

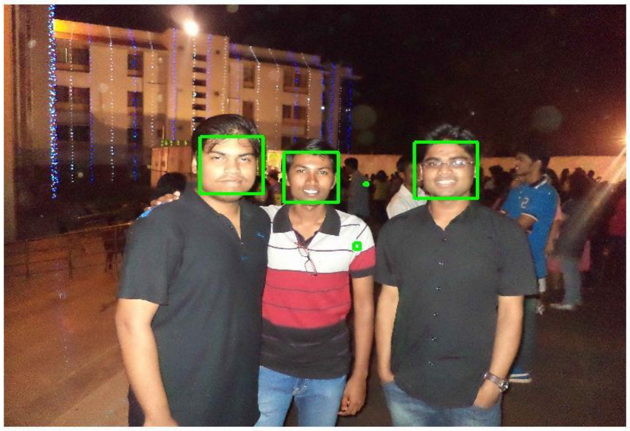

Figure 4.8 Detecting face from upper body

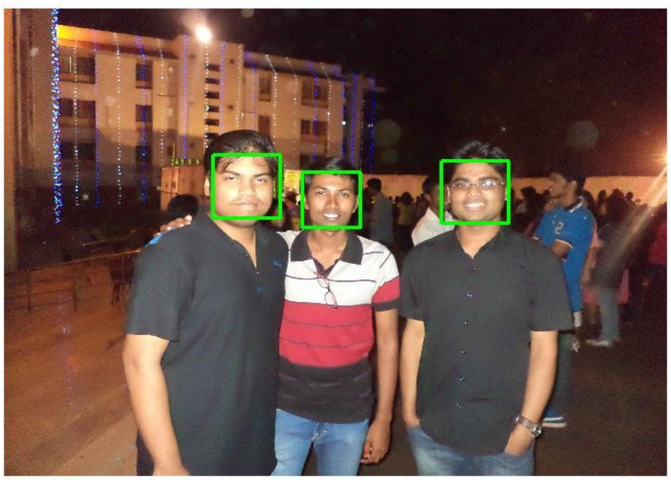

Figure 4.9 Remove non-faces

Chapter 5

Conclusion

We have exhibited a methodology for face detection which minimizes reckoning time while accomplishing high recognition precision. This calculation works just for frontal faces and can't be utilized for detection of side faces. Expecting the positive assessment set is a reliable representation of genuine appearances the current 16-stage locator is sensible to the point that in the ballpark of 97% of all (frontal upright) faces is distinguished. Without any bringing down of the edge this figure is diminished to approximately 95% which is still better than average. Likewise the measure of false positives is diminished and accordingly better execution attained. Experiments indicate that regularly more than 50 % percent of the false positives surviving the current 16-stage locator are moderately little. That is, their sizes are between 1 and 4 times the base determination of the indicator sub-window .

We removed the non faces in the image and only the faces are kept.

Bibliography

[1] Paul Viola, Michael J. Jones, Robust Real-Time Face Detection, International Journal of Computer Vision 57(2), 2004,USA .

[2] Paul Viola, Michael J. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features , Conference on Computer vision and Pattern recognition 2001,USA .

[3] Paul Viola, Michael J. Jones, Fast Multi-view Face Detection, Mitsubishi Electric Research Laboratories, TR2003-096, August 2003.

[4] Digital Image Processing Book, second edition, by Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins .

[5] Handbook of Face Recognition, by Stan Z. Li, Anil K. Jain